CLAIMS

1. A method of integrating multiple gene databases containing gene expression data, gene annotation data, and sample data, the method comprising:

defining new classes and associations for the gene expression data, gene annotation data, and sample data contained in at least one of the multiple gene databases;

grouping gene fragments into at least two groups; and

annotating gene fragments,

so as to permit a database analytical engine to analyze the gene expression data, gene annotation data contained in the multiple gene databases.

2. The method of integrating multiple gene databases containing gene expression data, gene annotation data, and sample data according to claim 1, wherein the at least two groups are known and not known.

3. The method of integrating multiple gene databases containing gene expression data, gene annotation data, and sample data according to claim 1, wherein the annotating if gene fragments comprises automatic annotation of predefined fields.

4. The method of integrating multiple gene databases containing gene expression data, gene annotation data, and sample data according to claim 1, wherein the annotating if gene fragments comprises manual annotation of individual gene fragments.

5. The method of integrating multiple gene databases containing gene expression data, gene annotation data, and sample data according to claim 2, wherein the grouping of gene fragments into at least two groups comprises name-based matching.

6. The method of integrating multiple gene databases containing gene expression data, gene annotation data, and sample data according to claim 2, wherein the grouping of gene fragments into at least two groups comprises sequence-based matching.

7. The method of integrating multiple gene databases containing gene expression data, gene annotation data, and sample data according to claim 2, wherein the grouping of gene fragments into at least two groups comprises manual data curation.

8. A computer program product comprising a computer-usable medium having computer-readable program code embodied thereon relating to multiple gene databases containing gene expression data, gene annotation data, and sample data, the computer program product comprising computer-readable program code for effecting the following steps within a computing system:

defining new classes and associations for the gene expression data, gene annotation data, and sample data contained in at least one of the multiple gene databases;

grouping gene fragments into at least two groups; and

annotating gene fragments,

so as to permit a database analytical engine to analyze the gene expression data, gene annotation data contained in the multiple gene databases.

9. The computer program product comprising a computer-usable medium having computer-readable program code embodied thereon relating to multiple gene databases containing gene expression data, gene annotation data, and sample data according to claim 8, wherein the at least two groups are known and not known.

10. The computer program product comprising a computer-usable medium having computer-readable program code embodied thereon relating to multiple gene databases containing gene expression data, gene annotation data, and sample data according to claim 8, wherein the annotating if gene fragments comprises automatic annotation of predefined fields.

11. The computer program product comprising a computer-usable medium having computer-readable program code embodied thereon relating to multiple gene databases containing gene expression data, gene annotation data, and sample data according to claim 8, wherein the annotating if gene fragments comprises manual annotation of individual gene fragments.

12. The computer program product comprising a computer-usable medium having computer-readable program code embodied thereon relating to multiple gene databases containing gene expression data, gene annotation data, and sample data according to claim 9, wherein the grouping of gene fragments into at least two groups comprises name-based matching.

13. The computer program product comprising a computer-usable medium having computer-readable program code embodied thereon relating to multiple gene databases containing gene expression data, gene annotation data, and sample data according to claim 9, wherein the grouping of gene fragments into at least two groups comprises sequence-based matching.

14. The computer program product comprising a computer-usable medium having computer-readable program code embodied thereon relating to multiple gene databases containing gene expression data, gene annotation data, and sample data according to claim 9, wherein the grouping of gene fragments into at least two groups comprises manual data curation.